Expert Systems with Applications 40 (2013) 2941–2946

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis

Sangjae Lee [a,*], Wu Sung Choi [b]

[a] College of Business Administration, Sejong University, 98 Kunja-dong, Kwangjin-gu, Seoul 143-747, Republic of Korea
[b] Seoul Credit Rating & Information Inc., 281-1/2 Sangsu-dong, Mapo-gu, Seoul 121-828, Republic of Korea

ARTICLE INFO

Keywords:
Bankruptcy prediction
Back-propagation neural network (BNN)
Multivariate discriminate analysis (MDA)
A multi-industry investigation

ABSTRACT

The accurate prediction of corporate bankruptcy for the firms in different industries is of a great concern to investors and creditors, as the reduction of creditors' risk and a considerable amount of saving for an industry economy can be possible. This paper presents a multi-industry investigation of the bankruptcy of Korean companies using back-propagation neural network (BNN). The industries include construction, retail, and manufacturing. The study intends to suggest the industry specific model to predict bankruptcy by selecting appropriate independent variables. The prediction accuracy of BNN is compared to that of multivariate discriminant analysis.

The results indicate that prediction using industry sample outperforms the prediction using the entire sample which is not classified according to industry by 6–12%. The prediction accuracy of bankruptcy using BNN is greater than that of MDA. The study suggests insights for the practical industry model for bankruptcy prediction.

1. Introduction

Prediction of corporate bankruptcy is of a great concern to investors/creditors, borrowing firms, and governments. As a result of the collapse of Enron, many voices have called for a revolution of existing bankruptcy warning systems to detect or prevent bankruptcy problems in real time. Bankruptcy can happen to any organizations because the business environment is increasingly undergoing uncertainty and competition these days. The reduction of creditors' risk and a considerable amount of saving for an economy can be possible from even a slight improvement with respect to assessing credit risk. An improvement in accuracy of even a fraction of a percent in scoring models to estimate the probability of default leads to enormous future savings for the credit industry (West, Dellana, & Qian, 2005). Assessment of bankruptcy offers invaluable information by which governments, investors, shareholders and the management can make their financial decisions in order to prevent possible losses. The study of bankruptcy provides an early warning signal and detects areas of weaknesses. Accurate bankruptcy prediction usually leads to many benefits such as cost reduction in credit analysis, better monitoring, and an increased debt collection rate.

A number of publications have pursued this subject and extending conventional models for prediction during the past 50 years. The number of bankruptcy prediction models has grown enormously due to the growing availability of data and the development of improved econometrical techniques during the 1980s and 1990s. Most of this work has been influentially led by a small number of early papers (e.g., Altman, 1968; Ohlson, 1980; Zavgren, 1985) on US quoted companies. The methods for bankruptcy prediction can be grouped in two categories: statistical and artificial intelligence models. The first group consists of Logit, multivariate discriminant analysis, etc. The tool first applied to bankruptcy prediction was the univariate data analysis proposed by Beaver (1966), which was followed by the multi-variate discriminant analysis and regressions (Ohlson, 1980). The second group includes neural networks (Chauhan, Ravi, & Chandra, 2009; Cho, Kim, & Bae, 2009; Pendharkar, 2005; Tseng & Hu, 2010), genetic algorithms (Etemadi, Rostamy, & Dehkordi, 2009; Lensberg, Eilifsen, & McKee, 2006), and support vector machine (Min & Lee, 2005; Yang, You, & Guoli Ji, 2011), and case based reasoning (Cho Hong & Ha, 2010). While some of these models show high predictive accuracy levels, the absence of bankruptcy theory makes attempts to establish a generally accepted model for bankruptcy prediction unsuccessful.

Although the discriminant analysis and linear regression model have become the most commonly used in bankruptcy prediction, their inherent drawbacks of statistical assumptions such as linearity, normality and independence among variables have constrained both applications. Recent trends in the development of artificial intelligence have brought forth new alternatives in solving nonlinear problems. The expert system, fuzzy logic, and neural networks are a great help to a manager in predicting bankruptcy making

* Corresponding author.
E-mail addresses: sangjae@sejong.ac.kr (S. Lee), karkasi@naver.com (W.S. Choi).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.12.009

decisions. Neural networks have many different topologies for problem dissimilarities. Among them, back-propagation is the most well known and commonly used, categorized as one of the supervised learning models. It draws the mapping function between the input and output from the provided data set. The back propagation neural network (BPN) usually contains one input layer, one or two hidden layer(s) and one output layer. Each layer of a neural network structure has several units and output units of a layer are input units of its next layer. The purpose of back-propagation training is to produce the weight of each edge, in order to minimize the squared error sum between the actual value and the predicted value.

Previous studies on neural network applications for bankruptcy prediction have been targeting single industry or not investigated the industry difference in bankruptcy prediction. For example, He and Kamath (2005) evaluated the effectiveness of two successful bankruptcy models by Ohlson (1980) and Shumway (2001) with the help of a mixed industry sample in discriminating between bankrupt and non bankrupt firms from an individual industry the equipment & machinery manufacturing (EMM) industry. Dewaelheyns and Van Hulle (2006) suggested that models involving both bankruptcy variables defined at subsidiary level and at group level provide a substantially better fit and classification performance. These studies did not examine the difference in industries in terms of independent variables, prediction accuracy, and practically usable models. It is still elusive whether generic prediction models are still successful in predicting individual industry. Further, although there exist a number of studies on using neural networks on bankruptcy prediction, the studies on using multi-industry data and developing models for multi-industry are almost rare. This paper intends to fill this void. This paper focuses on the different back propagation neural network (BNN) models for construction, retail, and manufacturing industries. This study intends to show the optimal tested for each industry by using the bankruptcy data of Korean companies. The study includes the comparison of predictive accuracy with multivariate discriminant analysis (MDA) and shows the implications of difference in prediction models and results.

2. Theoretical background

The credit risk analysis was pioneered by Beaver (1966). The author suggested cutoff threshold values for financial ratio variables in terms of profitability, liquidity, and solvency in order to classify them into two groups. The earliest studies about bankruptcy prediction were adopting the statistical approaches upon empirical data. Altman (1968) developed a statistical linear model and computed an individual firm's discriminant score to estimate the likelihood of bankruptcy. Altman used a combination of five financial ratios out of original list of 22 ratios. Ohlson (1980) utilized the logistic regression with a sigmoid function into the bankruptcy prediction problem. The logistic score, an outcome of logistic regression, is immediately interpretable into a statistical probability. Platt, Platt, and Pedersen (1994) examined the effect of deflation as a means of removing the temporal bias in data used to build a bankruptcy prediction model from the oil and gas producing industry. Clancy and Zhao (1999) suggested a bank failure prediction model based on banks' operations profile as financial intermediaries. Nam and Jinn (2000) investigated the predictive model of business failure using the sample of listed companies of a variety of industries that went bankrupt during the period from 1997 to 1998 when deep recession driven by the IMF crisis started in Korea. Neophytou and Molinero (2004) applied ordinal multidimensional scaling (MDS) to the corporate failure prediction problem, which relies on relations of order and does not suffer from extreme observations problems. Al-Attar, Hussain, and Zuo (2008) posited that abnormal accruals haven a small but significant degree of explanatory power while the explanatory power of abnormal accruals for future cash flows decreases at higher levels of bankruptcy risk. Leng, Feroz, Cao, and Davalos (2011) investigated the long-term performance and failure risk of firms cited in the Securities and Exchange Commission's (SEC) Accounting and Auditing Enforcement Releases (AAERs).

Researchers have applied artificial intelligence (AI) and data mining techniques to bankruptcy prediction. AI techniques are useful in finding out a non-linear relationship between variables. Data mining techniques are very actively utilized for searching an unknown meaning or non-linear pattern in a massive data set. Among these techniques, artificial neural networks have been applied in the context of bankruptcy prediction.

After neural networks are firstly used in bankruptcy literature in 1990, many researcher have applied neural networks in bankruptcy prediction. The use of neural networks has become popular due to the ability to learn and adapt from the data set, the ability to capture nonlinear relationships between variables, and the lack of need-to know functional forms a priori (e.g., Altman, Marco, & Varetto, 1994; Chen, Huang, & Lin, 2009; Sharda & Wilson, 1996; Tam & Kiang, 1992). Neural networks are good at capturing the mapping relationship among variables, especially nonlinear ones; however, they cannot explain the causal relationship among variables, which constrains its application to managerial problems.

3. Variables selection

In bankruptcy prediction, the main concern of interest is to construct the prediction model representing the relationship between the bankruptcy and financial ratios and then deploy the model to identify the high risk of failure in the future. A large number of features are usually included so that the training data is not enough to cover the decision space, which is represented as the curse of dimensionality. Feature selection represents the problem by excluding unimportant, redundant and correlated features in order to increase the accuracy and simplicity of classification model, reducing the computational effort, and enhancing the use of models.

The representative features for bankruptcy prediction can be presented as follows:

- Growth: Net sales growth rate, total asset growth rate, visible asset growth rate
- Profitability: Net profit/average assets, net profit/average shareholder's equity, income before taxes/average assets, interest income/total operating income, non-interest expenses/total operating income
- Productivity: Trading securities/total assets, fixed assets/fixed liabilities, net interest income/average assets, net on balance sheet position/total shareholder's equity
- Liquidity: Liquid assets/total assets, total loans/total deposits
- Asset quality: Permanent assets/total assets, total loans/total assets, loans under follow-up/total loans, specific provision/total loans, specific provision/loans under follow-up, shareholder's equity/total assets, shareholder's equity/total loans

4. Methods

In order to extract the variables that are importantly related to bankruptcy in each industry, *t*-test and correlation analysis are used in establishing the industry prediction model for BNN and MDA. The hit ratio is compared between BNN and MDA using *t*-test. KIS database is used for extracting the sample of the study.

The companies that became officially bankrupt from January 1, 2000 to December 31 2009 which became delisted from Korea Stock Exchange, comprised the bankrupt companies in the sample. The other companies which are not bankrupt were selected from the listed companies in the same industry which bankrupt companies belong to and they have the similar size of asset with bankrupt companies. The total number of companies is 229 including 91 bankrupt companies. 100 financial ratios are comprising each record for the companies. Table 1 suggests the number of companies in the Learning and test sample.

The number of test sample is adjusted to reflect the ration between bankrupt and non-bankrupt companies in the learning sample. The learning of BNN was repeated for 20,000 iterations. The BNN has a three layer architecture which has 6, 6, 1 nodes in the input, hidden and output layer. The logistic regression function was used as a transfer function.

5. Results

5.1. Selection of input variables

Using *t*-test, the significant variables which differ across bankrupt group and non-bankrupt group were selected from 100 financial ratios available in each record. The *t*-test to extract the important variables for being related to bankruptcy results in 46, 40, and 58 significant variables for construction, retail, and manufacturing industry at 0.05 significance level. For these selected variables, the correlation analysis was performed to select the variables which is most related to bankruptcy. Table 2 suggests the number of significant variables from t-test and correlation analysis.

The results of *t*-test indicate that the significance of group of variables is different among different industries. For construction industry, the growth, productivity and liquidity related variables are important for the prediction of bankruptcy. For retail industry, the stability and liquidity related variables are the major portion of significant variables. The growth, productivity, and stability related variables are crucial for the prediction of bankruptcy in manufacturing industry. The significant variables from the t-test are used in the correlation analysis with the dependent variable in order to select the final independent variables for BNN and MDA. From correlation analysis, one most correlated variables in each group is selected. These variables selected from five groups of variables can reflect various value characteristics of corporations which are either bankrupt or not. For retail industry, growth related variables are not included as they turn out to be insignificantly related to bankruptcy. Table 3 indicates the variables selected in this study.

5.2. Results

Neuroshell2 4.0 is used as the software for BNN. The number of independent variables for is 6, 4, 6 for BNN and MDA. Network topology describes the pattern of connectivity between nodes across a neural net. The approach in this study is based on a multilayer network topology. Nodes are layered with feedforward connections from lower to higher layers, which is called as feedforward network. Logistics function is adopted as an activation function.

$$OUT = \frac{1}{1 + e^{-\mathrm{NET}}}$$

where NET is the sum of weighted inputs and *OUT* is the final output of a neuron in the output layer. *OUT* is bounded by (0, 1). The function translates the activation value into an output value.

For the development of the knowledge-based network, this study uses BNN algorithm due to its simplicity and its usually good performance. The back-propagation algorithm adopts an iterative gradient descent method to adjust the weight associated with the connection between nodes to minimize the mean square error between the actual output of a multilayer feedforward net and the desired output. Each processing element produces an activation value by summing the weighted inputs. The activity level of that processing node is determined from the sum by an activation function in each node. If there is a difference between the output generated by the network and the known target value, the resulting error term is propagated back through the network, using a gradient- or steepest descent heuristic to minimize the error term by adjusting the connection weights. These weights are varied over time to improve performance based on current results and this provides a neural net with a degree of adaptability and robustness by compensating for minor variabilities in characteristics of the node.

The neural networks may have a three layer architecture consisting of input, hidden, and output layers. The use of a number of hidden nodes may lead to fitting the training set too well and decrease the generalization capabilities, although it increases training performance. Additional experiments show no significant differences in training and predictive performance among the configurations that have different number of hidden nodes ranging from two to eleven. In this study, the number of hidden nodes is set to be equal to the number of input nodes that moderately satisfy the objective of both training performance and generalization. Output layer has one node and each node in one layer is connected in the forward direction to every node in the next layer.

The system environments considered as affecting ISB2C controls constitute input variables. Input variables are accounting ratios. The output variable is bankruptcy. For construction industry, six inputs (x_1, x_2, x_3, x_4, x_5, x_6) are passed to each computational node *i* in the neural network along with six corresponding weights (w_{1i}, w_{2i}, w_{3i}, w_{4i}, w_{5i}, w_{6i}) where w_{ji} is the weight from node *j* to node *i*. The activation function produces the level of excitation by comparing the sum of these weighted inputs with the threshold value. This value is entered into the transfer function to derive the output from the node.

The cases are randomly split into two subsets, a training and test sample. Selecting appropriate training and test data sets is an important key decision for the successful implementation of BPN model. The training sample is used to build neural network models or regression equations and estimate their parameters.

To stop the training of BPN model, epochs since min. average error exceed a specified number is selected as the stop training criteria. The numbers that correspond to the stop training criteria for the training set is 20,000. The network will stop the training as soon as any one of the criteria is reached.

The average hit ratios in construction industry are 96.7 and 92.0 for learning and test sample. The average hit ratios in retail industry are 89.2 and 89.3 learning and test sample while the same ratios for manufacturing industry are 94.6 and 90.8. Table 4 indicates the predictions results for BNN.

For construction industry, the misclassification ratio for bankrupt companies in test sample is 12% which is larger than the one for non-bankrupt companies which is 6%. In a similar vein,

Table 1
Learning and test sample.

Sample		Construction	Retail	Manufacturing
Learning sample	Bankrupt	24	27	37
	Non-bankrupt	48	37	49
Test sample	Bankrupt	1	1	1
	Non-bankrupt	2	2	1

Table 2
Number of significant variables from *t*-test.

Group of variables	Total number of variables	Construction		Retail		Manufacturing	
		Number	Percentage	Number	Percentage	Number	Percentage
Growth	23	12	52.17	7	30.43	19	82.61
Productivity	35	23	65.71	16	45.71	19	54.29
Stability	22	8	36.36	12	54.55	15	68.18
Liquidity	5	4	80.00	3	60.00	2	40.00
Asset quality	15	2	13.33	2	13.33	3	20.00

Table 3
The variables selected in this study.

Groups	General model	Industry specific model		
		Construction	Retail	Manufacturing
Growth	Total asset growth rate	Total asset growth rate	–	Net sales growth rate
Productivity	Operating income to total asset Machinery equipment ratio per capital	Net profit ratio before income tax expense Net profit ratio before income tax expense per capita	Net profit ratio before income tax expense	Operating income to total asset Total assets per capita
Stability	Retained earnings to total asset	Retained earnings to total asset	Retained earnings to total asset	Retained earnings to total asset
Liquidity	Operating cash flow to total liabilities	Operating cash flow to total liabilities	Operating cash flow to total liabilities	Operating cash flow to total liabilities
Asset quality	Operating capital turn-over rate	Receivables turn-over rate	Operating capital turn-over rate	Operating capital turn-over rate

Table 4
The prediction results of BNN.

Sample	Real state	Prediction state		Total
		Bankrupt	Non-bankrupt	
(a) Construction industry				
Training sample	Bankrupt	575 (95.8%)	25 (4.2%)	600
	Non-bankrupt	34 (2.9%)	1141 (97.1%)	1175
Test sample	Bankrupt	22 (88.0%)	3 (12.0%)	25
	Non-bankrupt	3 (6.0%)	47 (94.0%)	50
Average hit ratios for training sample		96.68%		
Average hit ratios for test sample		92.00%		
(b) Retail industry				
Training sample	Bankrupt	672 (89.0%)	84 (11.0%)	756
	Non-bankrupt	103 (10.0%)	905 (90.0%)	1008
Test sample	Bankrupt	22 (79.0%)	6 (21.0%)	28
	Non-bankrupt	6 (7.0%)	78 (93.0%)	84
Training sample		89%		
Test sample		89%		
(c) Manufacturing industry				
Learning sample	Bankrupt	1292 (91.9%)	114 (8.1%)	1406
	Non-bankrupt	46 (3.0%)	1512 (97.1%)	1558
Test sample	Bankrupt	35 (92.1%)	3 (7.9%)	38
	Non-bankrupt	4 (10.5%)	34 (89.5%)	38
Training sample		94.60%		
Test sample		90.79%		

the misclassification ratio for bankrupt retail companies in test sample is 21% which is larger than the one for non-bankrupt companies which is 7%.

For manufacturing industry, however, the misclassification ratio for bankrupt companies in test sample is 7.9% which is lower than the one for non-bankrupt companies which is 10.5%. The overall misclassification ratios for bankrupt companies, however, are larger than the ones for non-bankrupt companies in the training and test sample.

Table 5 shows the prediction results for MDA in construction, retail, and manufacturing industry using each pair of learning and test sample.

MDA predicts bankruptcy in a similar hit ratio across three different industries. The misclassification ratio for bankrupt companies in test sample for construction industry (24.0%) is largely greater than the one for non-bankrupt companies (14.0%). Further, the misclassification ratios for bankrupt companies in test sample (31.8%) is far greater than the one for non-bankrupt companies (15.2%). This is also the case for manufacturing industry (26.3% vs. 10.5%). Table 6 indicates the summary of prediction accuracy for test sample in BNN and MDA. The paired *t*-test is used to examine whether BNN significantly outperforms MDA in average hit ratios in the test sample. Table 7 shows the results of *t*-test of the difference of hit ratios between BNN and MDA.

Table 5
The prediction results of MDA.

Sample	Real state	Prediction state		Total
		Bankrupt	Non-bankrupt	
(a) Construction industry				
Training sample	Bankrupt	422 (70.3%)	178 (29.7%)	600
	Non-bankrupt	31 (2.6%)	1144 (97.4%)	1175
Test sample	Bankrupt	19 (76.0%)	6 (24.0%)	25
	Non-bankrupt	7 (14.0%)	43 (86.0%)	50
Training sample			88.23%	
Test sample			82.67%	
(b) Retail industry				
Training sample	Bankrupt	563 (74%)	193 (26%)	756
	Non-bankrupt	112 (11.1%)	896 (88.9%)	1008
Test sample	Bankrupt	19 (68.2%)	9 (31.8%)	28
	Non-bankrupt	13 (15.2%)	71 (84.8%)	84
Training sample			82.71%	
Test sample			80.36%	
(C) Manufacturing industry				
Training sample	Bankrupt	1112 (79.1%)	294 (21.0%)	1406
	Non-bankrupt	215 (13.8%)	1343 (86.2%)	1558
Test sample	Bankrupt	28 (73.7%)	10 (26.3%)	38
	Non-bankrupt	4 (10.5%)	34 (89.5%)	38
Training sample			82.83%	
Test sample			81.58%	

Table 6
Prediction accuracy of BNN and MDA.

	BNN (%)	MDA
Construction industry	92.00	82.01
Retail industry	89.28	80.96
Manufacturing industry	90.79	81.58
Total sample	81.43	74.82

Table 7
Paired-t test for prediction ratios between BNN and MDA in test sample.

Industries	Average prediction ratio		Difference		t-Value	Significance (two tail)
	BNN (%)	MDA (%)	Average	Standard deviation		
Construction	92.00	82.01	9.990	4.083	2.447	0.022
Retail	89.28	80.96	8.327	3.686	2.259	0.032
Manufacturing	90.79	81.58	9.211	3.703	2.488	0.017

Table 7 indicates that the hit ratios of BNN are significantly greater than those of MDA in all three industries. Thus, BNN outperforms MDA in prediction accuracy in three industries. The learning and test sample can be changed to be randomly sampled on the condition that the proportions of bankrupt and non-bankrupt companies are 60% and 40% to reflect real prediction mode. The study further tests the prediction accuracy when the learning and test sample are randomly partitioned. Using these learning and test sample, BNN and MDA can be further applied in order to produce the hit ratios. It turns out that BNN again outperforms MDA in hit ratios.

To improve the readability of neural networks, the strength of relationship between each input and output variable can be represented by the following measure (Yoon, Guimaraes, & Swales, 1994):

$$RS_{ji} = \frac{\sum_{k=0}^{n}(W_{ki} * W_{jk})}{\sum_{i=0}^{m} ABS\left(\sum_{k=0}^{n}(W_{ki} * W_{jk})\right)_{ji}}$$

where w_{ki} indicates the weight between the kth hidden unit and the ith input unit, and W_{jk} means the weight between the jth output unit and the kth hidden unit, RS_{ji} is the relative strength between the ith input and jth output variable. This study uses RS_{ji} as the effect of independent variables on bankruptcy. This statistic indicates the strength of the relationship between the ith input and the jth output variable to the total strength of all of the input and output variables. The use of RS_{ji} provides the explanation facility for BNN model. RS_{ji} indicates the objective relative strength between variables before their values are given. Once the value of the effect is given, rules can be developed to display the value with interpretation.

Table 8 shows the relative strength of weights which represent the relative importance of links among input, hidden and output layer. The high absolute value of relative strength of each independent variables indicates the importance to determine bankruptcy. In retail industry, productivity, stability and liquidity which are represented by net profit ratio before income tax expense, retained earnings to total asset and operating cash flow to total liabilities, respectively, are more important factors to determine bankruptcy than in other industry. Asset quality in terms of operating capital turn-over rate is a more crucial factor to determine bankruptcy in manufacturing industry than retail industry. In construction industry, productivity and asset quality in terms of net profit ratio before income tax expense and receivables turn-over rate, respectively, are more critical factors to predict bankruptcy. Thus, growth, productivity and asset quality are important for construction and manufacturing industry, while stability and liquidity are crucial for retail industry. This shows that importance of net profit ratio, operating income, and turn-over rate of asset in construction and manufacturing industry; retained earnings and operating cash flow are critical in retail industry. These results are based on the different industry characteristics in terms of turn-over rate of fixed or liquid assets, proportion of fixed or liquid asset, and importance of cash flow and profitability.

Table 8
Relative strength of independent variables.

Class of variables	Independent variables	RS_{ji} on bankruptcy		
		Construction	Retail	Manufacturing
Growth	Total asset growth rate	0.13		
	Net sales growth rate			0.01
Productivity	Net profit ratio before income tax expense	−0.26	0.09	
	Net profit ratio before income tax expense per capita	0.33		
	Operating income to total asset			0.50
	Total assets per capita			0.19
Stability	Retained earnings to total asset	−0.03	0.69	0.05
Liquidity	Operating cash flow to total liabilities	0.15	0.16	0.08
Asset quality	Receivables turn-over rate	−0.10		
	Operating capital turn-over rate		0.06	0.17

6. Conclusions and implications

This study provides the multi-industry bankruptcy prediction model while previous studies are largely lacking in offering industry-specific model for bankruptcy prediction. This study intends to select the different group of independent variables to predict bankruptcy for construction, retail, and manufacturing industry. The study presented multi-industry prediction model using accounting-based measures as variables to predict bankruptcy and showed that financial ratios provide early warning signals. This model is better to reflect the industry characteristics in variable selection and to explain the different prediction results for each industry based on growth, productivity, stability, liquidity, and asset quality. In order to facilitate the application of BNN to the area of management, this study further provides relative strength of each independent variable for bankruptcy prediction, which may help improve interpretation of neural network learning and prediction. This will partially overcome the limitation that neural networks cannot explain the causal relationship amongst the variables.

The prediction accuracy is 6–12% greater for industry specific prediction model than the general model which has prediction accuracy of 81.43 for BNN and 74.82 for MDA. Further, it turns out that neural network models outperform MDA across three industries. This shows that neural network model better capture the nonlinear pattern between independent variables and bankruptcy than MDA. The prediction accuracy becomes greater for industry specific prediction model as the independent variables are specifically selected in each industry sample. Thus, it is necessary to build industry specific prediction model to produce more accurate and interpretable results.

References

Al-Attar, A., Hussain, S., & Zuo, L. Y. (2008). Earnings quality, bankruptcy risk and future cash flows. *Accounting and Business Research, 38*(1), 5–20.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*(4), 589–609.

Altman, E., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance, 18*, 505–529.

Beaver, W. H. (1966). Financial ratios as predictors of failure, empirical research in accounting: Selected studies. *Journal of Accounting Research, 4*(1), 71–111.

Chauhan, N., Ravi, V., & Chandra, D. K. (2009). Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks. *Expert Systems with Applications, 36*(4), 7659–7665.

Chen, H.-J., Huang, S. Y., & Lin, C.-S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. *Expert Systems with Applications, 36*(4), 7710–7720.

Cho, S., Hong, H., & Ha, B.-C. (2010). A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance; for bankruptcy prediction. *Expert Systems with Applications, 37*(4), 3482–3488.

Cho, S., Kim, J., & Bae, J. K. (2009). An integrative model with subject weight based on neural network learning for bankruptcy prediction. *Expert Systems with Applications, 36*(1), 403–410.

Clancy, D. K., & Zhao, R. (1999). A bank failure prediction model based on bank operations profile. *Asia-Pacific Journal of Accounting, 6*(2), 255–274.

Dewaelheyns, N., & Van Hulle, C. (2006). Corporate failure prediction modeling: Distorted by business groups' internal capital markets? *Journal of Business Finance & Accounting, 33*(5/6), 909–931.

Etemadi, H., Rostamy, A. A. A., & Dehkordi, H. F. (2009). A genetic programming model for bankruptcy prediction: Empirical evidence from Iran. *Expert Systems with Applications, 36*(2), 3199–3207. Part 2.

He, Y., & Kamath, R. (2005). Bankruptcy prediction of small firms in an individual industry with the help of mixed industry models. *Asia-Pacific Journal of Accounting & Economics, 12*(1), 19–36.

Leng, F., Feroz, E. H., Cao, Z., & Davalos, S. V. (2011). The long-term performance and failure risk of firms cited in the US SEC's accounting and auditing enforcement releases. *Journal of Business Finance & Accounting, 38*(7-8), 813–841.

Lensberg, T., Eilifsen, A., & McKee, T. E. (2006). Bankruptcy theory development and classification via genetic programming. *European Journal of Operational Research, 169*, 677–697.

Min, J. H., & Lee, Y.-C. (2005). Bankruptcy prediction using support vector machine (SVM) with optimal choice of kernel function parameters. *Expert Systems with Applications, 28*, 603–614.

Nam, J.-H., & Jinn, T. (2000). Bankruptcy prediction: Evidence from Korean listed companies during the IMF crisis. *Journal of International Financial Management & Accounting, 11*(3), 178–197.

Neophytou, E., & Molinero, C. M. (2004). Predicting corporate failure in the UK: A multidimensional scaling approach. *Journal of Business Finance & Accounting, 31*(5/6), 677–710.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131.

Pendharkar, P. C. (2005). A threshold-varying artificial neural network approach for classification and its application to bankruptcy prediction problem. *Computers & Operations Research, 32*, 2561–2582.

Platt, H. D., Platt, M. B., & Pedersen, J. G. (1994). Bankruptcy prediction with real variables. *Journal of Business Finance & Accounting, 21*(4), 491–510.

Sharda, R., & Wilson, R. L. (1996). Neural network experiments in business-failure forecasting: Predictive performance measurement issues. *International Journal of Computational Intelligence and Organizations, 1*(2), 107–117.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business, 74*, 101–124.

Tam, K. Y., & Kiang, M. Y. (1992). Managerial applications of neural networks: The case of bank failure predictions. *Management Science, 38*(7), 926–947.

Tseng, F.-M., & Hu, Y.-C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications, 37*(3), 1846–1853.

West, D., Dellana, S., & Qian, J. (2005). Neural network ensemble strategies for financial decision applications. *Computers & Operations Research, 32*, 2543–2559.

Yang, Z., You, W., & Guoli Ji, G. (2011). Using partial least squares and support vector machines for bankruptcy prediction. *Expert Systems with Applications, 38*(7), 8336–8342.

Yoon, Y., Guimaraes, T., & Swales, G. (1994). Integrating artificial neural networks with rule-based expert systems. *Decision Support Systems, 11*, 497–507.

Zavgren, C. V. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic analysis. *Journal of Business Finance & Accounting, 12*(1), 19–45.